



(U) Revealing Hidden Data at Scale Using AI: LANL's Terrier Project Delivers Permission-Aware AI for Faster Innovation and Discovery

Michael Ham

March 3rd, 2026

Data Infrastructure and Workflows

LA-UR-26-20041

Terrier is LANL's New Way to Discover Weapons Data

Natural Language Processing:
Utilizes specialized models to analyze and structure ingested documents.



Automated File Summarization:
A large language model generates file summaries



RAG Chat: An intelligent chat interface powered by extracted data allows dynamic user engagement.



Domain-Specific Filters: SME domain knowledge for locating datasets and context based information.



Terrier: a Unifying Way to Discover Data

LANL's Weapons Research Services is leading the way on digital transformation



Terrier AI powered data discovery platform



Data Engineers help get data into the right place (Poster 7)



CONDOR

CONDOR is a repository for official records (Poster 8)



National Ignition Facility (NIF; LLNL.gov)



NSDS Dataset Library

NATIONAL SECURITY DATA SOLUTION

NSDS is a repository for experimental and reference data (Poster 14)



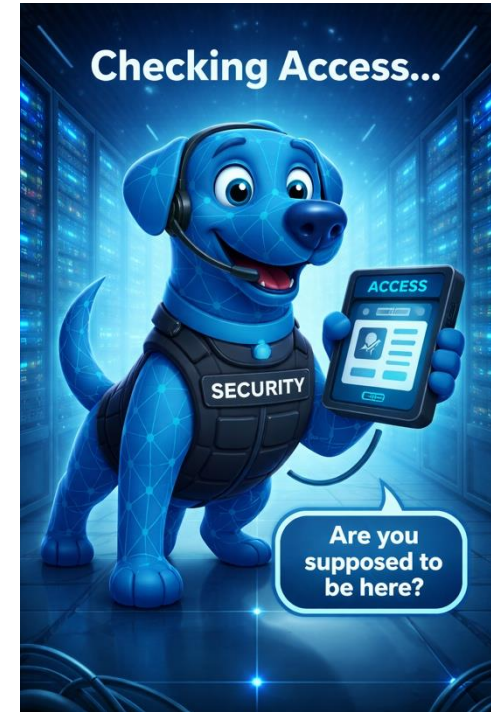
LANL supercomputers can pull data to perform mission work

Terrier's Number One Mission is Security and Enforcing Need to Know

- Terrier depends on robust data governance
 - Mission Data Stewardship Alliance (Midas)



- Terrier depends on WPDGC, PRIDE-DAG, DTSSG and other efforts to define FAIRified need-to-know access groups





Terrier's Ingestion Process is Constantly Updated and Enhanced

- Titan Technologies Compendia 
- Text extraction through machine learning algorithms
- Document summaries are automatically generated
- LLM's are easy to integrate as needed
- Natural Language Processing uses SME input and AI to link data through ontologies and semantic layers
- Dynamically link documents to need-to-know access groups during ingestion
- Terrier's modularity allows upgraded capabilities to be dropped into a production environment quickly
- **Future- data flows where it is needed automatically during ingestion**



Terrier: AI That Discovers Data For You



Search Preview View metadata AI Summary

When entering search terms, you are presented with possible options

In document view, you will see fewer results, but more detail for each result. The result with the gold header is your selected result (right). Click another result to change this.

In document view, you can view each document when you click on it from the left pane. You will see the full document (if you have access) and can search and view highlighted results from your search query.

Toggle highlights (from your search)

Advanced filtering combines 1) AI generated filters 2) SME generated filters

Filters you currently have selected (applying to your search)

Geopolitical Locations	Non-Geopolitical Locations	Military UAVs
<input type="checkbox"/> United States of America 181	<input type="checkbox"/> Earth 78	<input type="checkbox"/> Kronstadt Orion 1
<input type="checkbox"/> Nevada 106	<input checked="" type="checkbox"/> Savannah River Site 49	<input type="checkbox"/> MQ-9 Reaper 1
<input type="checkbox"/> California 101	<input type="checkbox"/> Bethel Valley Road 39	
<input type="checkbox"/> Federal Government of the U... 88	<input type="checkbox"/> Europe 33	
<input type="checkbox"/> Livermore 76	<input type="checkbox"/> North America 22	
<input type="checkbox"/> Las Vegas 73	<input type="checkbox"/> Asia 21	
<input type="checkbox"/> DIA Reports	<input type="checkbox"/> Mars 20	
<input type="checkbox"/> DOE	<input type="checkbox"/> Texas 18	
<input checked="" type="checkbox"/> DOE & National Laboratories	<input type="checkbox"/> Tennessee 16	
<input type="checkbox"/> DOG-Strategy-Documents	<input type="checkbox"/> Oak Ridge 10	
<input type="checkbox"/> DOS-Treaty-Documents	<input type="checkbox"/> Lawrence Livermore National... 63	
<input type="checkbox"/> DTRA Documents	<input type="checkbox"/> Sandia 61	
<input type="checkbox"/> Demo Documents	<input type="checkbox"/> Washington, D.C. 61	
<input type="checkbox"/> DoD1	<input type="checkbox"/> New Mexico 59	
<input type="checkbox"/> HALO	<input type="checkbox"/> Japan 52	
<input type="checkbox"/> NOAA_Docs	<input type="checkbox"/> Southern California 9	
<input type="checkbox"/> OB-Related Reports	<input type="checkbox"/> People's Republic of China 51	
<input type="checkbox"/> OONI Reports	<input type="checkbox"/> Alken 46	
<input type="checkbox"/> Proliferation-Issues	<input type="checkbox"/> New York City 42	
	<input type="checkbox"/> Albuquerque 40	
	<input type="checkbox"/> Northern California 7	

This is an unclassified example using locations to filter by. On the red, you can expect to see filters (tiles) related to nuclear weapons ontologies, such as nuclear tests and weapon types.

Numbers represent how many items fall under each category (expected results once checking the box)



• Terrier: AI that lets you chat with data

Ask your data questions

Terrier suggests topics and questions based on documents selected and your searches

Provides links to documents that answers use for reference

Easy to swap out LLMs for chat!

Venado integration coming soon

The screenshot displays the Terrier AI chat interface. At the top, a search bar contains the text "Search". Below it, a yellow banner indicates "6 Results (Individual/Aggregated)". A navigation bar includes options for "Map", "Tags", "Duplicates", "Chat", and "Ontology". A red note at the top left states: "Note: You currently cannot use AI chat for all documents in Terrier at once. Please limit to a specific topic or collection first before chatting." The main content area shows search results for "What are these documents about?". The results include sections for "Document Summary", "Abstract", "Recommendation Status", "Whale Sighting Surveys", "Whale Entanglement", "Species Recovery", and "Takeaways". At the bottom, there is a "References" section with a link to "[34017:0]". A chat input field at the very bottom contains the text "Ask the AI chat another question". Red arrows point from text annotations to various parts of the interface: one points to the "Chat" button, another to the "Individual/Aggregated" toggle, and a third to the "References" section.

Note: You currently cannot use AI chat for all documents in Terrier at once. Please limit to a specific topic or collection first before chatting.

6 Results (Individual/Aggregated)

Map Tags Duplicates Chat Ontology

What are these documents about?

Document Summary

Abstract

The provided documents cover a range of topics related to marine wildlife monitoring, conservation efforts, and safety recommendations. They include status reports on recommendations, whale sighting surveys, entanglement incidents, and species recovery plans.

Recommendation Status

One document is a status report detailing the completion and ongoing development of recommendations across categories like reporting, communications, training, and equipment. It focuses on observer safety and emergency preparedness. [34017:0]

Whale Sighting Surveys

Another document summarizes findings from the North Atlantic Right Whale Sighting Survey (NARWSS) and Right Whale Sighting Advisory System (RWSAS), aiming to reduce ship collisions with endangered whales. [34018:0]

Whale Entanglement

A third document details whale entanglement incidents along the U.S. West Coast in 2019, noting species affected and gear types involved. [34019:0]

Species Recovery

The final document outlines a 5-year action plan for the recovery of Central California Coast Coho Salmon, focusing on habitat restoration and hatchery improvements. [34020:0]

Takeaways

- The documents highlight ongoing efforts to monitor and protect marine wildlife populations.
- Conservation and recovery plans emphasize habitat restoration, reducing threats, and collaborative partnerships.

Keywords

Marine Wildlife Conservation Recovery Whales Salmon Entanglement Sighting Surveys Recommendations

References

> [34017:0]

Ask the AI chat another question

After asking the chat a question (in this case "What are these documents about?"), it will provide you with a summary

Click to initiate chat feature

Check "Individual" to chat only on the document you have selected, or "Aggregated" to chat on the entire set of search results

Terrier's Challenges

- Funding to launch
- Hardware is necessary, and cannot be faked
- High side network architecture
- Preventing hallucinations
- Getting the AI to perform the way we think it should



Terrier's Modular Framework is Constantly Evolving

- Integration with WRS products and beyond!
- Internal Development:
 - Identifying improved vision models to read data
- Partnerships:
 - The AI4NS project: (U) Using AI to Parameterize Legacy Diagnostic Analysis Sets
 - J. Clayton et al. NNS & LANL joint project
 - Production assistance
- Integration with ESN - TBD

